DOCUMENT RESUME

ED 448 751 IR 020 514

AUTHOR	Wade, Vincent P.; Doherty, Paul
TITLE	A Meta-Data Driven Approach to Searching for Educational Resources in a Global Context.
PUB DATE	2000-11-00
NOTE	11p.; In: WebNet 2000 World Conference on the WWW and Internet Proceedings (San Antonio, TX, October 30-November 4th, 2000); see IR 020 507.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Comparative Analysis; Computer Uses in Education; *Educational Resources; Foreign Countries; *Information Retrieval; *Information Services; *Internet; Models; *Systems Development
IDENTIFIERS	*Educational Brokerage; Europe; Institute of Electrical and Electronics Engineers; *Metadata; Search Engines; XML

ABSTRACT

This paper presents the design of an Internet-enabled search service that supports educational resource discovery within an educational brokerage service. More specifically, it presents the design and implementation of a metadata-driven approach to implementing the distributed search and retrieval of Internet-based educational resources and compares its performance with current search services, e.g., Internet search engines. The paper first presents a business model, identifying the possible different actors (customer, access network provider, broker, educational service provider, value added service provider, and content provider) in the context of a global educational open market. The major components of education metadata as defined by the IEEE (Institute of Electrical & Electronics Engineers) are summarized, including general, lifecycle, meta-metadata, technical, educational, rights management, relation, and annotation components). An architecture that reflects this model is then outlined. In focusing on the role of the educational broker, the current state of the art with respect to a metadata driven approach to educational content description is examined. More specifically, the paper concentrates on the metadata-driven search and retrieval issues when designing, implementing and trailing such a brokerage search manager, including the metadata cache and mapping XML to the relational model. The paper concludes with experiences of operating such services in a pan-European trial as part of the GESTALT research project. (Contains 10 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 448 751

A Meta-data Driven Approach to

Searching for Educational Resources in a Global Context

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

G.H. Marks

Vincent P. Wade
Paul Doherty
Dept. of Computer Science
Trinity College,
Dublin 2, Ireland
Vincent.Wade@cs.tcd.ie
Paul.Doherty@cs.tcd.ie

Abstract: As more and more educational services (e.g. on-line courses) and learning resources become available on the Internet today, it becomes increasingly difficult to identify those services or resources, which are of genuine relevance to the learner. Current Internet search services generally return very large numbers of results and it is left to the learner to glean the relevance of these selections. With the increased globalisation of the educational marketplace, this problem will become even more difficult. This paper proposes the design of a meta-data driven search management system for educational resources. The paper describes the design of an on-line educational brokerage service, which assists learners in identifying appropriate educational resources with greater accuracy than today's Internet search engines. The paper describes a novel XML/RDBMS implementation of the search management system and outlines the experience operating such a system on the Internet today. The paper concludes by providing observations on current IEEE standardisation for educational metadata and educational search services.

Introduction

A major impact of the Internet and more particularly the World Wide Web has been to assist the 'Globalisation' of services and markets. This has been seen in many eBusiness based ventures, from electronic retail (e.g. amazon.com) to interactive service e.g. stock exchanges and electronic trading. Globalisation is also beginning to influence the educational community. Morrison points out that "the globalisation of economic activities is forcing all nations to establish wider access to learning using communications technologies to create an 'educational options map' involving the cost, scale, quality, relevance, portability, futurity, flexibility of an access to education" (Davis 99, Morrison 95). There are many challenges in an evolution toward global educational markets including cultural, economic, political, social and technical. Although it can be argued that such globalisation may not be appropriate across *all* educational domains, it's growth is certainly evident in tertiary level continuing education and in lifelong learning. From a technological perspective, many issues must be tackled e.g. standardisation between interfaces of educational systems and services, standardisation of description of educational resources and services to enable identification and enable appropriate choices, brokerage services to support potential learners in navigating the maze of WWW offerings and providing some guidance for choice. This paper focuses on services to support learners/tutors in identifying and choosing courses or learning objects (resources) in a global context. Current search approaches e.g. Internet search engines provide only basic search parameter matching and relevancy of retrieval is poor. The approach adopted in this paper is that of a supporting independent educational broker(s) who can provide prospective learners or tutors with an informed, independent view of many educational offerings/resources. The proposed search management system returns more precise search results and provides better precision based on the learners requirements. From the educational service provider's perspective, the brokerage service can provide a cost effective means of attracting learners from different geographical and cultural areas. For example on the internet today, 'knowledge portals' are beginning to appear where separate organisation(s) have been set up to market their on-line learning programmes and products from various educational providers.

IR020514

BEST COPY AVAILABLE

This paper presents the design of an Internet enabled search service, which support educational resource discovery within an educational brokerage service. More specifically it presents the design and implementation of a meta-data driven approach to implementing the distributed search and retrieval of Internet based educational resources and compares its performance with current search services e.g. Internet search engines. The search system, although trailed in a pilot educational brokerage service, could also be used across any set of federated educational providers who support explicit metadata descriptions of their resources. However, in this paper we focus on a separate brokerage service applied within the educational domain. The paper first presents a business model, identifying the possible different actors in the context of a global educational open market. The paper then outlines an architecture, which reflects this model. In focusing on the role of the educational broker, the paper examines the current state of the art with respect to a meta-data driven approach to educational content description. More specifically the paper concentrates on the meta-data driven search and retrieval issues when designing, implementing and trailing such a brokerage search manager. The paper concludes with its experiences with operating such services in a pan-European trial as part of the European Research project GESTALT (HREF 1).

A business model & architecture for educational brokerage

Ideally, in a global open market there would exist a (standards based) framework for the development of compatible, heterogeneous, scalable, and distributed educational systems globally distributed across institutions and organisations. Such a framework would allow individuals to discover the existence of desired educational resources and have these resources delivered remotely over established networking infrastructures via a managed learning environment. One approach to realising this framework would be to analyse the potential organisation types and actors (roles) and interaction between these actors (relationships), which would form such a global vision. Such a vision is sometime called a 'Business Model'.

The European Research Project GESTALT analysed the requirements and influences in offering WWW based commercial educational services and distilled this into the GESTALT business Model. The authors do not claim this is the only Broker model for Global Education Services, but it is one based on current commercial / Educational influences from the internet. Several other models are possible e.g. consortia of educational service providers combining to offer branded educational resources or learning programmes. The GESTALT Business Model describes the roles involved in an Educational process and the relationships between these roles. In figure 1, the boxes represent the roles and the lines with circled numbers represent the relationships between roles.

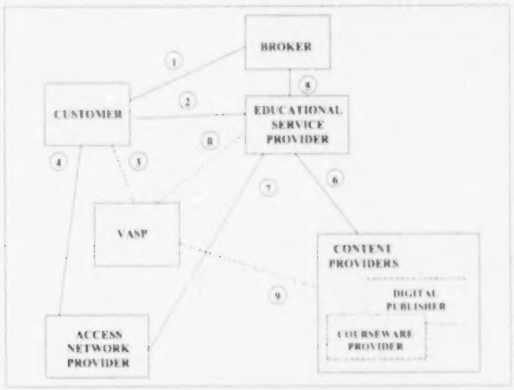

Figure 1: GESTALT Business Model

It is important to realise that there need not be a one-to-one correspondence between a role and an organisation. In fact in realising such a model, several roles can be played by a single organisation. The following is a brief description of the roles played in the GESTALT environment.

BEST COPY AVAILABLE

Roles within the GESTALT business model

Customer - at the heart of the GESTALT Business Model is the Customer. An individual or organization playing the role of customer in the GESTALT environment is the entity wishing to discover the existence and location of an educational resource and/or actually participating in a GESTALT educational session.

Access Network Provider – by its very nature the GESTALT environment is distributed with the various GESTALT roles being played by actors residing at different locations. The Access Network Provider is the organization providing the physical links and network services required to link together the various actors playing GESTALT roles.

Broker – an electronic broker can be defined as an on-line entity that supplies specialized services or products, or information about such services/products to customers wishing to discover or purchase them. In the GESTALT environment the service/product will be the Courses and Resources (either as discrete components or packaged into an educational session) or learning programmes offered by an educational service provider and the customer will be the entity functioning in the GESTALT role of Customer wishing to discover the existence of an educational resource or set of educational resources meeting customer defined search parameters.

Educational Service Providers – the educational service provider is the entity providing educational services to the GESTALT role of customer. The responsibilities for this role can be classified into the following three categories:

- Learning Object Management - providing structured storage and management of learning objects
- Learning Object Delivery - providing an on-line environment for the delivery of learning content in the form of learning objects to the GESTALT customer.
- Student Tracking - providing administration facilities to handle the registering and course progress of educational sessions followed by customers registered with the educational service provider.

Value Added Service Provider (VASP) – the GESTALT role of value added service provider will provide User Profile Services to the GESTALT roles needing such services. A user profile provides a common point for the structured consolidated storage of information about an individual and all of the educational and vocational experience obtained by the individual.

Content Provider - the individual or organization actually producing and making available the educational content comprising learning objects.

In the business model where a consortium of educational service providers, the brokerage role can be managed by an independent company or be offer virtually by one of the consortium members (usually prime partner in the consortium). In this paper we focus on the Search Management aspects of the broker. We analyse the requirements and interactions with the Learner and Educational Service Provider. In the model, the Learner (customer) uses a brokerage site to identify educational offerings, which are suited to her. The brokerage service provides retrieval of learning programmes/objects more relevant for the learner's requirements. The search manager presented in this paper uses a meta-data driven approach to resources identification and is based on emergent IEEE standards in educational meta-data. The next section outlines the current state of the art in educational resource modelling and then presents the design and implementation of a standards based educational brokerage service.

Using Meta-data for Educational Search Management

Currently there are several international initiatives in the area of learning technology and learning object (resource) modelling. Principal among these is the IEEE Learning Technologies Standards Committee (LTSC) (HREF 2), which has formed a working group to create a standard for Learning Object Metadata (LOM) (HREF 3). This work has been heavily influenced by several research initiatives both in Europe and US. The Instructional Management Systems (IMS) (HREF 4) project is a US initiative, originally driven by EDUCAUSE, which incorporated 600 educational institutions across the USA. A second significant research initiative was ARAIDNE (HREF 5), a research project funded under the EU Telematics Programme. Both of these initiatives have influenced the early design of the LOM. Subsequently other research projects and organisations have also contributed to the standard e.g. GESTALT, moving LOM to its current relatively stable

state. Another significant research initiative in the area of metadata has been the 'Dublin Core Metadata Initiative' (HREF 6) whose approach is 'aimed at achieving generalised interoperability across interest/subject domains and is focused on resource discovery'. Because of Dublin Core's 'cross subject' approach (i.e. applicable to many problem domains) their first set of meta data attributes were both limited in number and structurally simple. However, the group is currently working on a draft specifically for the Educational domain.

The major components of such Education metadata as defined by the IEEE LOM are:

- **General** - Groups all context independent features plus the semantic descriptors for the resource e.g. description, author, domain
- **Lifecycle** - Groups the features linked to the lifecycle of the resource e.g. version, creator of the course, publisher, title, ...
- **Meta-metadata** - Groups the features of the description itself (rather than those of the resource being described) e.g. creator of the metadata
- **Technical** - Groups the technical features of the resource e.g. format of the resource, size of the resource
- **Educational** - Groups the educational and pedagogic features of the resource e.g. pedagogical type, semantic density
- **Rights Management** - Groups the features that depend on the kind of use envisaged for the resource
- **Relation** - Groups features of the resource that link it to other resources e.g. identifier of related resource
- **Annotation** - Allows for comments on the educational use of the resource e.g. person, description, date

A detailed breakdown of the LOM structure can be found on the IEEE web site (HREF3), a brief overview of the standard reveals: There are 101 elements in version LOM version 2.5 and all elements are optional; There is no required subset of metadata attributes; The Dublin Core metatdata elements are all represented in the LOM structure but are grouped differently; The LOM is structured into a hierarchy and its format is expressed in XML. On first reading the standard, the meaning of each element is not always clear, nor are the values required for certain elements are not always clear; the model is extensible for future evolution. The search manager described in this paper utilises LOM v2.5 and conclusions as to it usefulness and operation for resource discovery is discussed in the final sections of the paper.

Architecture

The purpose of the Broker is to provide a reliable gateway through which potential students can be referred to quality Courses and Resources. The students must have confidence that the Broker will point them at dependable courses and resources therefore unlike search engines only content from a relatively few quality assessed providers need be offered. Since Courses and Resources may exist on other networks the Broker cannot be limited to the WWW e.g. Z39.50 provides a valuable mechanism for accessing the wealth of information available in existing library systems. A major component of the Broker is a search facility similar to that offered by today's search engines. Whilst current search engines do an admirable job of indexing the majority of the WWW and offer solid text search functionality over HTML, their results often lack precision. To provide a more accurate search facility the Broker is required to prompt the user with domain specific information, which will help in the specification of a more precise search. To personalise searches for users the Broker is required to store profile information e.g. a user's preferred spoken language. Facilities for the previewing, and ordering of courses and Resources are also of use to perspective students and are therefore Broker requirements.

The key element of any Educational Search facility is the data about the courses and resources that are offered, in other words metadata about the courses and resources. HTML the ubiquitous data format for the web, is very successful at marking up text to be displayed on a browser. However HTML does not facilitate marking up with domain specific information. The W3C (HREF 7) proposes that XML be used to address this limitation and since domain specific information is a key requirement of the Broker, XML was adopted to facilitate its provision. This domain specific data provides a common schema through which authors and users can better communicate.

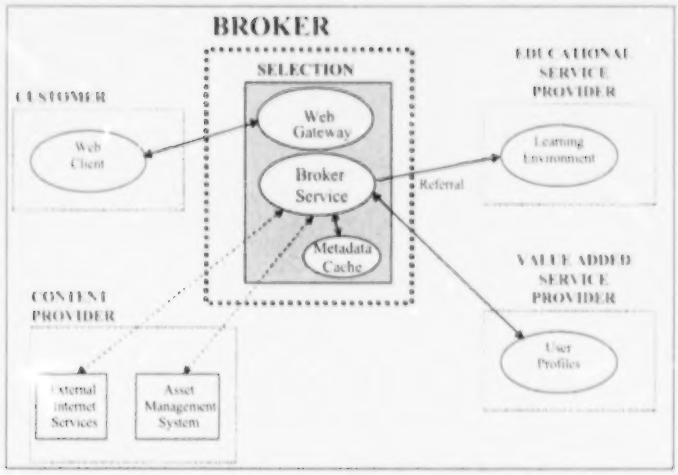

Figure 2: Implementation Architecture for the Educational Broker

The Broker depicted in Figure 2 uses a standard three-tiered architecture. The First Tier is a Web Client for use by the potential learner, the middle tier contains the application logic of the Broker Service, and the third is a metadata cache of educational courses and resources. The other components in Figure 2 are components from the GESTALT business model, which are external to the Broker. A thin **Web Client**, provides a Graphical User Interface for interaction with the rest of the Broker Service. The Web Client is used for the entry of registration, search, course preview, and course order requests. These requests make calls to correspondingly named Server Manager Objects in the **Broker Service**. The first of these Objects is a Registration Manager which governs access to the system, standard login and logout, as well as caching of user profiles. The Registration Manager can if required interact with a **Value Added Service Provider** to access centrally stored user profile information. The second Object is an Education Search Manager which accesses a cache of Course and Resource metadata, and provides search facilities over that cache. The metadata is cached in order to provide adequate performance. This cache of metadata is obtained by downloading metadata at regular intervals from different **Content Providers**. Backend Agents in the Broker Service are used for this gathering process and these Agents are similar to the Web Crawler's used by today's Search Engines. However, content providers are not restricted to Web Sites as they can be a database, a Z39.50 based library system, or any other system which stores Course or Resource metadata. The Agent is responsible for any protocol conversion issues and converting the data in the Content Provider to XML. Where the Course and Resource data to be retrieved by an Agent is not a format the can readily be converted to XML, a metadata conversion tool can be used along side the Agent to generate appropriate XML metadata (Kenny99). The final Object in the Broker Service is an Order Manager which interfaces with the **Educational Service Provider** to facilitate student referral and course ordering.

In order to satisfy the requirements for the ordered and previewing of courses, the Broker defines protocol independent interfaces. Further interfaces are defined for the connection of backend Agents. The main parameter in most of these interfaces is the XML metadata, other parameters include standard URLs for the identification of courses and resources, and a free format text string for user names. Because most of the application specific information is contained in the XML the defined interfaces have relatively few parameters and should NOT change significantly with time.

Trial broker implementation

The Broker was implemented for the GESTALT trial using a number of technologies including XML, CORBA, LDAP, and a Relational Database. The Broker is designed as a distributed system and therefore CORBA,

BEST COPY AVAILABLE

Common Object Request Broker Architecture, was selected to facilitate communication between the various components. In a large configuration where the course metadata needs to be allocated across many machines, the CORBA Trader is used to select appropriate physical instances of the Broker. The Broker was successfully demonstrated to international external reviewers and the European commission in November 1999, and has been operating over a period of four months across several institutions.

At the outset of the project it was unclear which (educational) metadata schema would be used. However, if the Broker could store any metadata schema then it could be used for any metadata driven application not just Education. The Education Search Service is therefore able to store and search any XML data schema and is not restricted to the LOM. The cache keeps sufficient of the original XML data to be able to provide the Web Client with the following information:

1. Which DTDs schemas are available for searching. Each DTD schema can be thought of a single application domain, due to time and other constraints no attempt was made to map between different metadata schemas or different versions of a single schema.
2. Which elements exist in a given DTD and therefore which elements exists for a given domain.
3. Which data values exist for a given element, useful where the number of data values is small e.g. a list of all Universities.

None of these features are available in today's Internet search engines. These facilities assist the user in defining a very precise search in accordance with the selected schema, with today's search engines a user does not received any such domain specific assistance.

The Broker offers three types of searches of increasing precision. The first is equivalent to that offered by many of today's search engines, the second makes use of the metadata structure to focus the search into a section of the XML, and the third adds user profile information. The three types of searches are referred to as Internet search, metadata search, and compound metadata search. The metadata search offers the user a list of element names from the selected metadata schema e.g. Title, Description, Author etc. The user is then able to perform searches on these elements e.g. Author CONTAINS "Smith" or Keyword CONTAINS "C++". The compound metadata search automatically adds user profile information, which can come from a centralised profile management system, refining the search the user entered e.g. the condition "Language = Italian" would be automatically added if the users preferred language was Italian. This compound metadata search will be more precise then the metadata search. Standard Boolean operators are supported e.g. AND, as are standard comparison operators e.g. Greater Than. The Web Client can offer these searches either as a command line interface or as a Graphical User Interface which prompts the users with the different domains, element names, and element values.

JAVA as selected as the programming language for the Broker Management Objects and a Relational Database was used for the metadata cache. By using JAVA and JDBC, the search facility is easily ported to different operating systems and relational DBMSs e.g. the Education Search Service was ported from NT and Microsoft Access to UNIX and Oracle with minimal code changes. A JAVA Database Object was coded which made use of JAVA reflection to create SQL commands so that persistent images of objects could be stored in a Relational table. All persistent Classes were derived from the Database Classes and inherited the ability to create themselves as a table in the Relational Database. Instantiated objects of these classes had the ability to read, write, delete, and update themselves to/from rows in the Relevant Relational table. This technique made the creation of new persistent objects and changes to the elements in persistent objects very easy, it also made the creation of a new database very straightforward as all tables were created automatically from the Java code.

The metadata cache / mapping XML to the relational model

The major goals for searching the cache were acceptable performance, flexibility of searching with as much domain specific knowledge as possible, and the ability to ignore the structure of the XML and to search all the element values individually. It was known that XML could be stored in a relational DBMS and that RDBMS's offer a powerful query languages. Further the Authors are of the opinion that existing technology should be used were ever possible and therefore the decision was taken to attempt to store the XML in a RDBMS and map the inbound user query to SQL.

Sample Conversion of XML into an SQL table

This example illustrates the major ideas involved in storing XML in a Relational database and how the querying of that model works. Consider the following XML sample:

```
<Student>
        <Name>
                <First>Paula<\First>
                <Last>Jones<\Last>
        <\Name>
        <Mentor>
                <Name>
                        <First>Sarah<\First>
                        <Last>Murphy<\Last>
                <\Name>
        <\Mentor>
<\Student>
```

This sample is a piece of Student information about someone called "Paula Jones", who has a Mentor called " Sarah Murphy". There are a number of note worthy points in the XML structure. Firstly there are element names e.g. Student, Name, First etc which must be presented in any structure which stores the XML. Secondly there are different levels in the XML. "Student" is at the highest level or level one, whereas Mentor is at low/deeper level, level two. Note that "First" is both at level three and at level four and that the difference between the two levels is significant. At level two "First" implies the Students name and at level four "First" implies the Mentor's name. Another way to look at this difference is to consider element names and expanded element names, both have a element name of "First" and have expanded element names of "Student.Name.First" and "Student.Mentor.Name.First" respectively.

One possible way to represent the XML in the relational database is to store the XML document locations in one SQL table, expanded element names in a second SQL table, and the element values from the XML files in a third. Thus we have the three tables:

XML Info Table

ID	Location
1	www.cs.tcd.ie/patient.xml

Expanded Element Info

ID	Element Name
1	Student
2	Student.Name
3	Student.Name.First
4	Student.Name.Last
5	Student.Mentor
6	Student.Mentor.Name
7	Student.Mentor.Name.First
8	Student.Mentor.Name.Last

Element Value Info Table

XML ID	Expanded Element ID	Value
1	3	Paul
1	4	Jones
1	7	Sarah
1	8	Murphy

Each time a new XML document is processed a single new entry representing the document location is added to the XML Info table and for a given DTD schema each element is represented uniquely by one row of the Expanded Element Info table. The possibility that two DTD schemas can use the same top-level name has been ignored. (This can be solved by adding a DTD ID to the table and modifying the query language to specify the required DTD(s).) For each data value in the XML fragment a new row is added to the Element Value Info Table.

Querying the cache

An example of a search that could be entered into any of today's search engines is:

Paula AND Jones

The Search Manager maps such a query to the following SQL statements:

```
SELECT DISTINCT   XMLInfo.Location
FROM              ElementValueInfo , XMLInfo
WHERE             ElementValueInfo.xmlID = XMLInfo.ID
          AND
                  xmlID IN (SELECT DISTINCT xmlID FROM ElementValueInfo
                  WHERE dataValue  = "Paula" )
          AND     xmlID IN (SELECT DISTINCT xmlID FROM ElementValueInfo
                  WHERE dataValue  =  "Jones" )
```

An example of a search qualified with element names is:

Student.Name.First = "Paula" AND Student.Name.Last = "Jones"

This is translated into the following three SQL statements:

```
SELECT DISTINCT elementID1 from ExpandedElementInfo
Where elementName = "Student.Name.First"

SELECT DISTINCT elementID2 from ExpandedElementInfo
Where elementName =  "Student.Name.Last"

SELECT DISTINCT   xmlID
FROM              ElementValueInfo
WHERE
                  xmlID IN (SELECT DISTINCT xmlID FROM ElementValueInfo
                  WHERE (Expanded Element ID  =  elementID1  AND  value =
                  "Paula") )
          AND     xmlID IN (SELECT DISTINCT xmlID FROM ElementValueInfo
                  WHERE (Expanded Element ID  =  elementID2  AND  value  =
                  "Jones"))
```

Other Query Functionality

The query language available to the user includes the following functionality:

1) Wild cards in the element names. Therefore it is possible to write searches such as:
   %First = "Paula" AND %Last = "Jones"
   This query will look for the first element name in the DTD ending in "First" and the first element name ending in "Last" and then perform the metadata search as previously explained. Thus this search returns identical values to that returned in the previous example.
2) Partial matches in the data values. Thus the user can enter:
   %First CONTAINS "Paul" AND %Last CONTAINS "Jone"
   Again this search returns identical results to that returned in the previous example.
3) The following Boolean operators are supported: AND, OR
4) The following comparison operations are supported: <> = < > <= >= CONTAINS
5) All results are returned in XML format.
6) The user can request any or all of the elements from the DTD schema and the Search Manager reconstruct either the entire XML document or fragments of the XML document. Again the user can use wild cards to indicate which elements are to be returned.

Results of the XML to relational mapping

Storing XML metadata in a RDBMS yielded mixed results. One the positive side the Broker is able to store any XML document no matter what the DTD schema, which is very flexible and therefore desirable. The inbound user query is mapped to SQL in order to perform the search which saved much development effort. It also

enables the middle tier to provide a powerful non DTD specific interface to the Web Client which in turn enables the Web Client to offer an interface based upon the domain (DTD) that the users is interested in not just a LOM specific interface. Further a Relational database is a secure repository, which can be accessible by other applications. On the negative side standard SQL text searching is limited and needs to be augmented to provide such functionality as stemming, a Thesaurus, etc. This could be achieved using something like Oracles ConText product or by alternatively by developing similar functional as a front end to the database. Due to time constraints neither of these was attempted during the project. The trail produced adequate performance results for a database of 137 metadata documents, more development and testing is being done before the performance issue can be said to be fully satisfied for a large-scale repository. There are some fundamental mismatches between XML and the relational model these are:

1. XML has no limit on element length, Relational databases do for elements which can be indexed. For the trail an maximum element length was assumed.
2. XML is hierarchical and permits a element to repeat. Relational databases do not permit repeating elements and their main structure is a table not a hierarchy. The XML must be represented in a number of SQL tables and the resulting queries are complex as they must span these tables.
3. XML only supports the data type text it does not support such simple types as integer, date, currency etc. This leads to problems when comparing numeric and date elements e.g. "3" will compare greater than "100" which is clearly not what is expected. The only way to solve this problem is to parse the XML and attempt to recognise the type of a given element. This was not done as part of our trial and it is not expected that this would be an easy problem to solve in a generic fashion.

Observations from the trial

These observations come from the authors experience of developing the Broker software and a comparison against their expectations and previous software development experience. The observations are also as a result of the operation of the broker in the GESTALT project. From a software engineering stand point it was noticed that the functionality of CORBA and XML overlap in that they can both be used to describe the messages passed between different parts of the system. In one respect this can be seen as an advantage as most of the application specific information is in the metadata and this is the data most likely to change, leaving the more stable parameters in the IDL bindings, which should not change significantly over time. From another perspective, however, it would have been possible to implement the broker with say JAVA Remote Method Invocation and XML based messages.

As part of the demonstration different European partners kindly contributed approximately 137 LOM format metadata files on different programming and networking courses. During the trial a number of cases were demonstrated showing metadata searches returning fewer matches than Internet searches. For example suppose a student was looking for a "C++" programming course. If an Internet search was done for the string "C++", it would return thirty matches. If a metadata search was performed using the Keyword element i.e. keyword = "C++" 5 matches would be returned. Finally a compound metadata search could be performed for an Italian-speaking student and would return 2 matches. The principal behind these searches intuitively makes sense, since if the keyword element is set to "C++" the course is likely to be about "C++", whereas "C++" can be present in other parts of the metadata but the course would NOT necessarily be about "C++".

The LOM metadata was not filled in a complete fashion, people filled in different elements in the LOM, and the same elements differently. Also authors reported that it took some considerable time to understand the schema, and stated that they would have appreciated a LOM specific tool to assist them. A further difficulty was that much of the LOM was of little use for resource discovery since much of it is used for the delivery of a resource. Finally casual user's performing a search did not find the LOM structure easy to navigate. This experience points to some "profile" (subset) of the LOM being agreed for resource discovery. The Dublin Core initiative on Education metadata was not complete at the time of the trial and it would be most interesting to contrast these results with a trail of Dublin Core.

The Broker is highly dependant on the DTD schema, as GESTALT's experience with LOM shows. As far as resource discovery, for a mass audience of casual users, a straightforward DTD schema will make it easy for course authors to create the metadata and easy for students to understand and navigate whilst searching. A complex structure will make it difficult for authors to create the metadata unless they are given a specific tool

and even then they will have difficulty if they are required to enter too many fields or do not understand what values a given field requires. Further a complex structure will make it hard for users to query unless some subset of the metadata is identified. It is not possible to define precisely what is meant by a straightforward DTD schema, some guidelines are:

1. A few required elements and attributes, let say less than thirty.
2. A clear definition of each element with examples, including its data type (integer, date, etc).
3. The use of attribute lists (pick lists) or fields with a limited number of values, were possible e.g. the names of all the colleges in Ireland.

Conclusions

The trail found that a metadata driven Resource Discovery Service is possible and more importantly that it can provide a more accurate search facility than today's search engines given a straightforward DTD schema (see previous paragraph for a definition of "straightforward"). However course authors must put in extra work to generate the metadata according to a standard schema(s). The LOM is a good beginning for the Educational domain however it is possible that a subset of the LOM should be defined for Resource discovery and that other Educational metadata schemas will emerge e.g. Dublin Core. In order to gain acceptance these schemas should have a straightforward design for the benefit of the course authors who must create the metadata and perspective students who must search the metadata. A new project funded under the IST European Research Program EASEL, is continuing the metadata approach for content searching and management. EASEL is building on the metadata specification and is extending these to support adaptive content and retrieval

A relational database can be designed to store any XML document. However querying that database is limited from the perspective of performance, and is certainly limited from the perspective of standard text searching feature such as stemming, Thesaurus etc. The authors are currently researching improvements to the initial design to overcome these limitations for very large-scale metadata repositories.

There are no fundamental design issues with respect to the Broker, which are outside what can be achieved by software engineering, with the proviso that no work was done to map between different metadata schemas, since only one Educational Schema was available to us. However a significant amount of engineering and other work still needs to be done to make a metadata driven approach a reality.

Acknowledgement

The authors offer thanks all their colleagues from the GESTALT project Richard Crawford, Yves Nicol, Willie Donnelly, Paul Foster, R Smith, C Lambrinoudakis, M Konstantopoulos, J S Darzentas, and Kevin Reilly. Special thanks to Benjamin Jung of Trinity College for his ideas in relationally modelling XML structures.

References

Davis, Niki (1999). 'The Globalisaion of Education through Teacher Educations with New Technologies: A View Informed by Research', *Educational Technology Review*, Autumn/Winder 1999, No. 12., pp 8-12 Publication of AACE.

Morrisson, T (1995). 'Global transformation and the search for new educational design', *International Journal of Lifelong Education*, 14,2, pp 188-213

Kenny, S. (1999). 'Identification and Retrieval of WWW-based Courseware – *A Metadata Approach*', MSc Thesis, Department of Computer Science, Trinity College Dublin. 1999

HREF 1. GESTALT D0201 Joint Requirements Specification http://www.fdgroup.co.uk/gestalt/d201v2.zip

HREF 2. LTSC, IEEE Learning Technologies Standards Committee http://grouper.ieee.org/groups/ltsc/

HREF 3. LOM, IEEE Learning Objects Metadata Version 2.5 http://grouper.ieee.org/groups/ltsc/ltsc_docs/wgc/LOMdoc2_5.doc

HREF 4. IMS, Instructional management Systems http://www.imsproject.org/

HREF 5. ARIADNE - Educational Metadata Recommendation Summary, Working Document, Version 2.0, April 1998 http://ariadne.unil.ch/

HREF 6. Dublin Core Metadata Initiative http://www.purl.org/dc/

HREF 7. W3C, World Wide Web Consortium http://www.w3.org/